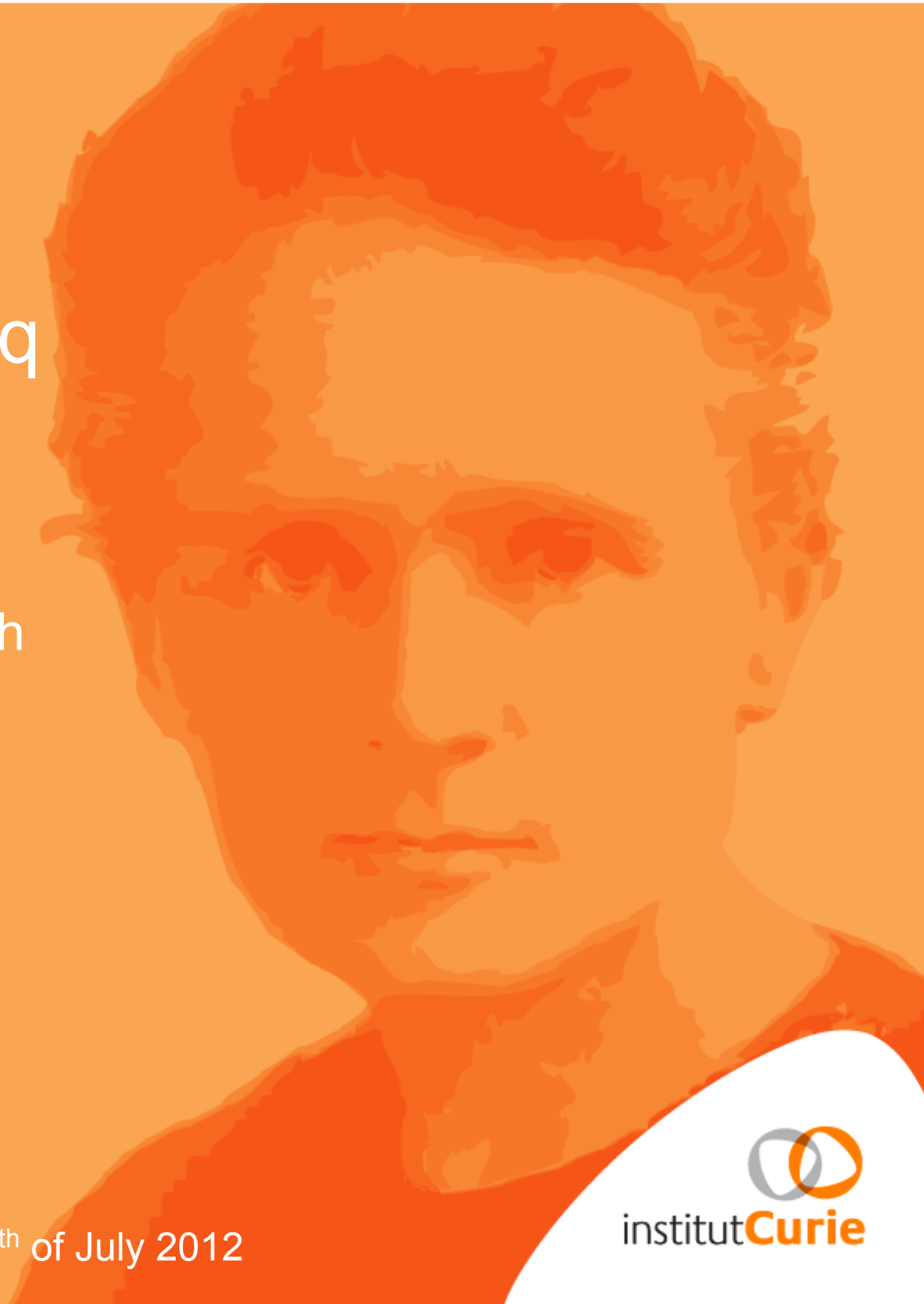# Nebula :
## A public web-server for advanced ChIP-seq data analysis

**Alban Lermine**
U900 INSERM - Mines Paris Tech
Institut Curie

**Galaxy Community Conference 2012** - 26th of July 2012

institut**Curie**

## 1. Overview – Galaxy at Institut Curie

- Galaxy is installed since my GCC2011 return (June 2011)

- 3 instances are installed:

 . One public ChIP-seq specific server - http://nebula.curie.fr

 . One local production server - http://galaxy.curie.fr

 . One local development server - http://galaxy-dev.curie.fr

- Each instance is connected to a job scheduler (PBS-Torque) on a different queue to dispatch jobs on our Cluster

- Nebula is a classic Galaxy server; Dev & Prod servers have following features:

 . LDAP authentication

 . No data upload through the network allowed

 . All jobs are submitted to the cluster "as the current user"

 . Each tool allows users to choose where they want to write their results

- Almost 30 new tools have been integrated since 1 year

institut**Curie**

## 2. Nebula

### 2.1. Nebula's main goal

- ChIP-seq analysis accessible for biologists



- Identify and annotate transcription factor bindings sites starting from sequencing raw reads

- INSERM Workshop: "ChIP-seq data analysis for Biologists"

institut**Curie**

## 2. Nebula

## 2.2. Technologies and softwares used in Nebula

Nebula is composed of 23 tools:

- Published tools:

    . Bowtie, Samtools, MACS, FindPeaks, ChIPmunk, Fastqc, BEDtools

- Homemade tools (Perl, R & Bash):

    . Get peak distribution around TSS,

    . Extract central regions of peaks,

    . Genomic annotation of peaks,

    . Gene annotation using peaks,

    . Get peak height distribution

institut**Curie**

## 3. Nebula's workflow

# Let's go to the Nebula's homepa

**institutCurie**

## 4. Availability



# http://nebula.curie.fr

*Genome analysis*

# Nebula – a web-server for advanced ChIP-Seq data analysis

Valentina Boeva[1,2,3]*, Alban Lermine[1,2,3], Camille Barette[1,2,3], Christel Guillouf[1,4] and Emmanuel Barillot[1,2,3]

[1]Institut Curie, Paris, F-75248 France, [2]INSERM, U900, Paris, F-75248 France, [3]Mines ParisTech, Fontainebleau, F-77300 France, [4]INSERM, U830, Paris, F-75248 France

- In Press !

institut**Curie**

# 5. Perspectives

Nebula:

- Nebula will be used for training at the French Bioinformatics School (Roscoff, Jan 2013)

- Integration to the main toolshed (need to adapt code)

- Integration of new tools to handle other types of ChIP-seq data (histone modifications)
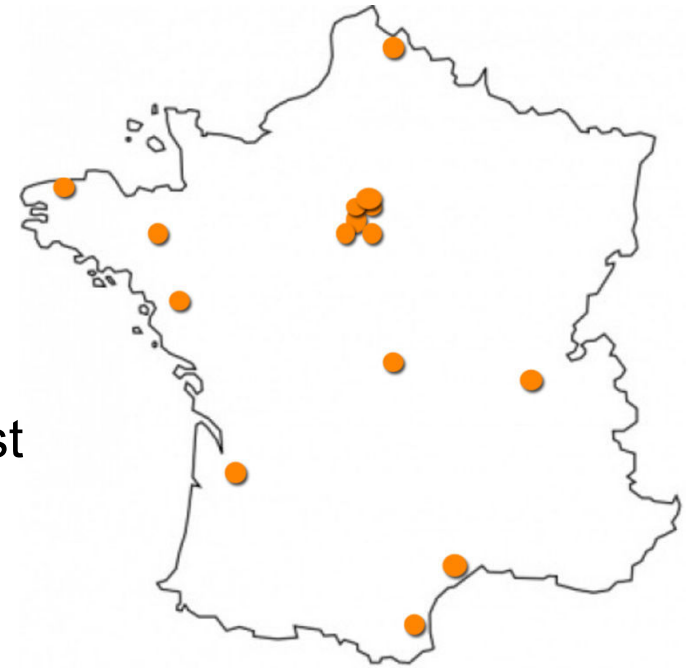
Galaxy:
- Nebula-like servers for RNA-seq (**Séverine Lair**), Exome-seq (**Bruno Zeitouni**), sRNA-seq (**Nicolas Servant & Chongjian Chen**)

- Nebula-like server for diagnostic data from IonTorrent (**Bruno Zeitouni**)

institut**Curie**

## 6. A word about French Galaxy community

- Thanks to the Galaxy Tour de France, we identified Galaxy users all over the country

- We share Galaxy knowledge and tools within the **RENABI ApliBio** consortium ("Ile de France" scale)

- With the help of the Galaxy team, especially **Dave Clements**, we will create a Galaxy french diffusion list

- With **Olivier Inizan** (URGI/INRA Versailles) and **Sophie Gallina** (GEPV/Université Lille 1), we plan to organize the first GFCC in early 2013

institut**Curie**

# Nebula - A **Galaxy** based web-server for advanced **ChIP-seq** data analysis

*V. Boeva, A. Lermine, C. Barette, C. Guillouf, E. Barillot*

## Special thanks to :



**Valentina Boeva**
*Research Engineer*
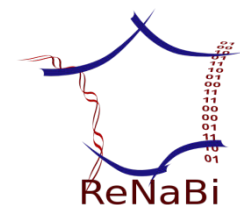U900 INSERM



**Camille Barette**
*System Engineer*
Institut Curie



**Christel Guillouf**
*Research associate*
U528 INSERM



**Emmanuel Barillot**
*Director of Bioinformatics*
U900 INSERM

 institut**Curie**

 **Inserm** Instituts thématiques · Institut national de la santé et de la recherche médicale

 MINES PARIS ParisTech

 Galaxy

 ReNaBi

institut**Curie**

## Great thanks to the Galaxy Team !

# Thanks for your attention !

institut**Curie**